Kristin “Krissa” Stewart

July 1st, 2021

LIM-503

Linked Data: The Future of Library Access

As librarians, one of the critical purposes of our career is to share and spread information openly. With the advent of the world wide web, this goal has seemed closer in hand than ever before. Yet, the spread of misinformation is just as large as the spread of factual information. As library professionals, we want to ensure that the data is accurate and easy to find for our patrons. Yet, going between different types of libraries, or even different library branches, can lead to confusion and frustration. If there is a standardized way to save the data libraries gather, and thus a standardized way to access that data, it would be easier for patrons and staff to access it. That is where the concept of Linked Data comes in. If all data is connected, there is no reason to memorize a thousand different categories to find the same book. There is no need to travel to another library just to see if they have something the last one you were at did not have. Simply put, Linked Data is an essential tool in making the research and location process easier for everyone involved. You are no longer wasting time traveling to find something that might not even be where you wish it were.

Linked Data is a relatively new concept in the realm of technology, which has been transforming rapidly since the invention of the home computer. What Linked Data is, essentially, is a structured way to post data onto the internet to try towards a fuller integration of the data in your system. Linked Data connects to all the data around it in a structured way, making it more useful in semantic queries (Linked Data 2015). It can also help draw inferences from the search, including related terms in the results and those searched for by a patron. This cuts down on time, thinking up synonyms and aids in finding more resources a patron may not have thought of before.

Linked Data also allows libraries, in particular, to enhance their web presence through search engine results (Schilling, 2012). By linking the data within library websites to larger systems and in a more structured way, it increases the chances that the library website will show up on the first page of a Google search result. The closer something is to the top of a Google search result page; the more likely the patron is to click on it. For example, suppose someone is searching for a specific book. In that case, the local public library can carefully structure the data so that their page about that novel will appear right below the Amazon and Barnes & Noble listings. A potential patron may then click on that link, then see a page on a book they are interested in or a program they may enjoy and be more willing to spend time in the library and participate in programs.

Linking the data also allowed for a more logical and straightforward way of searching. Without Linked Data, you can only search for what you know exists, and you are more limited by the words you use. If you use the wrong synonym, you fall into the problem that what you are looking for may not show up. But by linking the synonyms together in a structured Linked Data system, you will see much more related content because the synonyms are not considered separate from each other. Instead, the words or phrases are connected by their similarities. That way, fewer things get left out of a patron's search due to a minor issue.

Eventually, the goal of Linked Data is to connect everything into a Semantic Web. The idea is to make the web more accessible to computers and to allow computers a measure of AI that would allow them to think about a subject and make connections that otherwise may have gone unnoticed or thought of. Of course, with any measure of AI, people do worry about the AI getting to power, but that would not be an issue here. This system is simply a more complicated and more intelligent search engine that can make the lives of library staff and patrons easier.

The first step libraries took in adopting the Linked Data system was categorizing everything in their collections using Resource Description Framework, or RDF (Schilling, 2012). This is a sentence structure used for metadata that includes a subject, a predicate, and an object. This is a standardized system used to communicate metadata on the web. By standardizing metadata communication, that metadata can be shared more easily between branches and between different library systems and even different types of libraries.

Libraries specifically apply Linked Data by using Bibliographic Framework Initiative, or BIBFRAME. The Library of Congress started this project to replace the commonly used MARC as the standard for encoding library-generated metadata. This system is specifically made for libraries and works with information that libraries work with every day, such as bibliographic and authority records. This system makes it easy to include the information libraries need in the Semantic Web, which patrons may not necessarily need or even think about if they do not cite something for a paper or article.

The internet is a great connector. By categorizing items with RDF and then programming it into order with the Linked Data system, we as librarians can share more information with more people, thus increasing the opportunities for discovery, learning, and knowledge.

This is not to say there are no problems with Linked Data. One particular concern is that not all libraries have the staff, time, or technology to complete what is required to participate in the Linked Data movement. For example, a small-town library may have excellent genealogical pieces but also be understaffed, have slow internet, and out of date technology. So, while they may be interested in taking part in the Linked Data movement, they simply do not have the means to. To help with this issue, grant funding may be in order, but it is many years off if it were to happen. Linked Data may be popular with libraries right now and may move to be the industry standard one day, but it is not the standard yet.

Another issue with Linked Data is that it depends on standardized vocabulary. There are very few punished lists of that vocabulary (Schilling, 2012). For a system that runs on the standardization of metadata, this is a severe problem. To truly be a unified Semantic Web, everyone needs to be using the same vocabulary lists. If they are not, it becomes much harder to link that data. Seeing that this is the entire point of Linked Data, this could quickly spiral out of control.

We also see issues with copyright and reliability (Schilling, 2012). On the hand of reliability, human beings are known to make mistakes regularly. Even if someone has the best intentions and their work is double-checked but multiple other staff members, errors can slip through. That could cause the data to link to the wrong pieces of metadata, confusing the entire system without anyone realizing it.

When looking at copyright, it is essential to remember that copyright law struggles to keep up with the digital age. For example, how do we handle the sudden increase in eBooks? Most companies have taken to selling someone the rights to download the book rather than owning the physical book in a digital form. Is this how copyright would be applied to Linked Data at large? Or would media companies come up with something entirely different? This is impossible to predict. If a company such as Disney gets involved, Linked Data may be heavily regulated and charged or even made illegal ultimately.

Even with these issues, the library system, in general, is moving into using Linked Data. For many librarians, it appears that the benefits of a Linked Data system outweighs the negative. As stated earlier, many librarians see the exchange and sharing of information as a far greater task than any other. Having a new way to share information in an easily accessible format, no less, is something that librarians dream of. The easier it is for patrons, the more satisfied they will be with the library and their learning. In the end, the more information there is and the easier it is to find, the better it is for many libraries' missions.

While Linked Data is ready and many librarians are moving in that direction, it still has a long way to go before it becomes the standard practice for libraries across the United States. "The technology is ready; it is now a matter of getting libraries and librarians ready as well" (Schilling, 2012). It is important to remember that change is frequently a slow-moving beast. Some libraries and librarians may be stuck in their ways. Some may not have the money or staff to make the necessary changes. Some universities may not have faculty experienced enough to teach the next generation of librarians what they need to know to work with Linked Data. Sometimes it may end up in a matter of poverty and class, where some libraries can afford the funding for up-to-date internet and technology, while others cannot. After all, we still live in a country where one's income can determine what public services are available. In some areas of the country, getting a good internet connection is still extremely difficult.

Yet, with patience and determination, it is likely that in 10 or 15 years, Linked Data will be the industry standard. Sooner, if those less fortunate libraries can receive the funds to take part in making the linked metadata. Until then, it is essential to learn, stay up to date, and become involved inside and outside their systems. If librarians can work together, they may even be able to teach and aid more librarians, bringing this standardized system to fruition.

References

Schilling, V. (2012, September 26). Transforming Library Metadata into Linked Library Data. Association for Library Collections &amp; Technical Services (ALCTS). https://www.ala.org/alcts/resources/org/cat/research/linked-data.

W3C. (2015). Linked Data. W3C. https://www.w3.org/standards/semanticweb/data.